

周报（2014.05.05-2014.05.11）

- 1、继续之前的用 `python` 从网页上抓取数据，尤其是执行完查询后产生的动态网页的内容。学习用 `beautifulSoup Python` 库从抓取的网页上执行查询。目前已经实现了给定一个关键词，抓取在搜索引擎中动态搜索到的子网页的所有内容。
- 2、学习 `Xquery` 的相关内容。当我们根据用户输入的多个数据源建立起 `ontology` 以后，如何进行查询推理以及整个数据的管理是重点。`Xquery` 是 `W3C` 的一套标准，用来在 `XML` 文档中提取信息，也包括以 `XML` 语言呈现的其他数据或数据库。`XQuery` 可用于提取网页的信息以便在网络服务中使用或者为了获取相关的信息和执行某个查询。目前先对 `XQuery` 有了个大概的了解，在 `W3C` 上学习了 `XQuery` 的基本内容，包括基本的语法和较高级的添加和选取的操作。

下周工作：

- 1.需要学习如何用 `XQuery` 查询 `XML` 和 `OWL` 推理或者查询，要知道如何用编程逻辑执行这些查询。